

# EUROPEAN PATENT OFFICE

## Patent Abstracts of Japan

PUBLICATION NUMBER : 03154551  
PUBLICATION DATE : 02-07-91

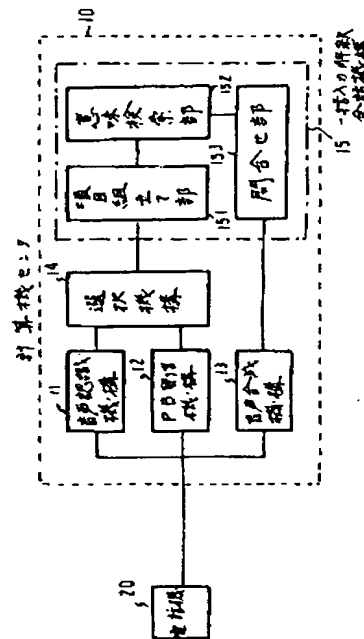
APPLICATION DATE : 13-11-89  
APPLICATION NUMBER : 01294358

APPLICANT : NIPPON TELEGR & TELEPH CORP  
<NTT>;

INVENTOR : NAKADA HISASHI;

INT.CL. : H04M 1/26 H04M 3/42

TITLE : TELEPHONE INPUT DEVICE



ABSTRACT : PURPOSE: To input plural items from a telephone set in the lump through the free combination of voice and a PB signal by using a selection device to obtain the result of processing of voice recognition mechanism against the voice input and to obtain the processed result of a PB reception device to the PB (pushbutton) signal.

CONSTITUTION: A signal inputted from a telephone set 20, that is, a voice or a PB signal is processed in parallel with a voice recognition device 11 and a PB reception device and the result of processing (code) is given to a selection device 14. The selection device 14 discriminates whether the input is a PB input or a voice input while receiving the result of processing of the voice recognition device 11 and the PB reception device 12, adopts the result of processing of the voice recognition device 11 and the voice input is discriminated and when the PB input is discriminated, the result of processing from the PB reception device 12 is adopted, then the result of processing of the selection device 14 is given to a simultaneous input decoding interactive device 15 and the plural items inputted simultaneously are correctly decoded, thus, the simultaneous input of plural items is attained.

COPYRIGHT: (C)1991,JPO&Japio

## ⑫ 公開特許公報(A) 平3-154551

⑤Int.Cl.<sup>5</sup>

識別記号

庁内整理番号

⑬公開 平成3年(1991)7月2日

H 04 M 1/26  
3/42

C

7190-5K  
7925-5K

審査請求 未請求 請求項の数 1 (全7頁)

⑭発明の名称 電話入力装置

⑰特 願 平1-294358

⑱出 願 平1(1989)11月13日

⑲発明者 勝 俣 雅 司 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内

⑳発明者 阪 上 晃 一 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内

㉑発明者 中 田 寿 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内

㉒出願人 日本電信電話株式会社 東京都千代田区内幸町1丁目1番6号

㉓代理人 弁理士 鈴木 誠

## 明 細 書

## 1. 発明の名称

電話入力装置

## 2. 特許請求の範囲

(1) 電話機から音声とプッシュボタン信号(以下、PB信号と称す)の組合わされた1または複数の項目を入力する装置であって、

電話機からの音声入力を認識し、対応するコードを出力する音声認識機構と、

電話機からのPB信号入力を検出し、対応するコードを出力するPB受信機構と、

前記音声認識機構とPB受信機構の出力を受けて、当該コードが音声入力かPB信号入力か判定し、音声入力と判定した時は音声認識機構の出力を選択し、PB信号入力と判定した時はPB受信機構の出力を選択する選択機構と、

前記選択機構の選択結果を予め定義した入力項目の構成情報にもとづいて解釈する解釈機構とからなることを特徴とする電話入力装置。

## 3. 発明の詳細な説明

## 〔産業上の利用分野〕

本発明は電話入力装置に係り、特に声音、プッシュボタン(PB)信号の両方を用いて電話から複数の項目を入力するのに好適な電話入力装置に関する。

## 〔従来の技術〕

電話サービスにおいて電話機から情報を入力する場合、入力方式には、計算機センタからのガイダンスで指定された項目を入力するガイダンス方式と、1回の入力で任意の複数(1も含む)の項目を入力する一括入力方式があり、また、入力信号については、声音とPB信号がある。

第4図ガイダンス方式の構成例を示したもので、10は計算機センタ、20は電話機である。計算機センタ10は電話機20からの入力を受け付けるもので、音声認識機構11、PB受信機構12、音声合成機構13、ガイダンス会話処理機構16よりなる。利用者は電話機20から音声またはPBにより情報を入力する。入力された信号は、計算機センタ10において音声認識機構11または

P B受信機構12でコードに変換され、ガイダンス会話処理機構16へ渡される。ガイダンス会話処理機構16は入力された信号を解釈し、利用者に対して音声合成機構13を用いてガイダンスを音声出力する。

第5図はガイダンス方式による利用者と計算機センタの会話シーケンス例で、(イ)は音声入力、(ロ)はP B入力、(ハ)は数字をP B入力では音声入力で行う例である。いずれも、1回の入力要求に対しては、音声かP Bのいずれかのみの入力であり、利用者は計算機センタから指定された項目を入力する。

第6図は一括入力方式の構成例で、音声による入力の場合を示したものである。計算機センタ10は音声認識機構11、音声合成機構13、一括入力会話機構15よりなり、一括入力解釈会話機構15はさらに項目組立て部151、意味検索部152及び問合せ部153よりなる。

電話機20から入力された音声は、計算機センタ10において、音声認識機構11で認識され、

しかしながら、ガイダンス方式には、インタラクション数が多くなることや、入力の必要がない項目にも「不要」などと答えなければならないといった問題がある。これに対し、一括入力方式では1回の入力で複数項目が入力可能であり、かかるガイダンス方式の場合の問題が解釈される。

一括入力方式において、1回の入力に音声とP Bを混在させると、例えば、数字入力は確実に入力できて操作が簡単なP B入力、数字以外の入力は覚え易い単語を用いた音声入力とし、これらを自由に組合せて一括入力することにより、効率的でマンマシンインタフェースのよい入力が期待できる。しかしながら、従来は1回の入力では音声、P B一方のみの入力を前提としているため、音声とP Bを自由に組合せて一括入力するようなマンマシンインタフェースのよい入力方式を実現できなかった。

本発明の目的は、音声とP Bを自由に組合せ、複数項目(1項目も含む)を電話機から一括入力する効率的でマンマシンインタフェースのよいデ

単語列(コード列)に変換されて、一括入力解釈会話機構15へ渡される。一括入力解釈会話機構15の各部では、この単語列に対して次のような処理を行い、複数項目を直しく解釈する。即ち、項目組立て部151は関連ある単語をまとめて複数の項目に組立てる。意味検索部152は項目をサービス毎に決められた入力項目に対応づける。問合せ部153は誤認識等の原因で生じた解釈不可部分を音声合成機構13を用いて利用者に関合わせる。なお、この一括入力解釈会話機構15については、例えば特願昭63-129349号に詳述されている。

第7図は第6図の一括入力方式による利用者と計算機センタの会話シーケンス例を示したものである。

〔発明が解決しようとする課題〕

ガイダンス方式では、1回に1項目の入力だけであるので、入力毎に音声、P Bを変えた入力が可能である。第5図(ハ)は、数字入力項目はP B、その他の入力項目は音声入力とした例である。

ータ入力を可能とする電話入力装置を提供することにある。

〔課題を解決するための手段〕

上記を達成するため、本発明の電話入力装置は、電話機からの音声入力を認識し、対応するコードを出力する音声認識機構と、電話機からのP B信号入力を検出し、対応するコードを出力するP B受信機構と、前記音声認識機構とP B受信機構の出力を受けて、当該コードが音声入力かP B信号入力が判定し、音声入力と判定した時は音声認識機構の出力を選択は、P B信号入力と判定した時はP B受信機構の出力を選択する選択機構と、前記選択機構の選択結果を予め定義した入力項目の構成情報にもづいて解釈する解釈機構からなることを特徴とする。

〔作 用〕

音声、P B信号の判定は、一般に音声周波数とP B信号周波数の帯域が異なることを利用して、周波数フィルタにより行われる。この場合、周波数フィルタにより、音声でP B受信機構が動作し

ないようにできても、音声認識機構がPB信号を誤って認識する危険性があり、音声とPB信号の2つの入力があったと誤判定されることがある。また、音声認識処理時間、PB検出処理時間は一般に異なるので、音声認識機構からの出力、PB受信機構からの出力順序が反転することがある。このような場合も、選択機構により、音声入力に対しては音声認識機構の処理結果、PB信号に対してはPB受信機構の処理結果が必ず得られるため、電話機から音声とPB信号の両者を使用して複数の項目を入力することができるようになる。

(実施例)

以下、本発明の一実施例について図面により説明する。

第1図は本発明の一実施例の構成図で、10は計算機センタ、20は電話機である。計算機センタ10は電話機20からの音声とPB入力を受け付けるもので、音声認識機構11、PB受信機構12、音声合成機構13、選択機構14及び一括入力解釈会話機構15よりなり、一括入力解釈会

音声認識機構11では、音声周波数とPB信号周波数の帯域が異なることを利用して音声を認識し、単語列(コード列)を出力する。この結果、音声入力(入力1、入力4)は、音声認識機構11からT<sub>v</sub>の処理遅延後、認識結果V<sub>1</sub>、V<sub>4</sub>として出力される。この音声認識機構11にはPB信号も入力されるため、該PB信号と標準パターンとの照合を行い、この結果が、ある閾値より近ければ、処理遅延T<sub>v</sub>後、音声認識結果が出力される。第3図の認識結果V<sub>2</sub>は、PB信号(入力2)を誤って認識したことを示している。閾値より離れていれば、認識対象語ではないと判定し何も出力されない(入力3、入力5、入力6に対しては音声認識機構11からの出力はない)。一方、PB受信機構12には、ある帯域の周波数を減衰または強調する周波数フィルタをPB信号受信部の前に置くことにより、音声入力を誤ってPB信号とすることを防ぎ、本当にPB信号が有った時のみ、該PB受信機構12が動作するようにすることが可能である(例えば特開昭61-78296

話機構15はさらに項目組立て部151、意味検索部152、問合せ部15よりなる。

電話機20から入力された信号(音声またはPB信号)は音声認識機構11、PB受信機構12で並行して処理され、処理結果(コード)が選択機構14へ渡される。選択機構14では、音声認識機構11とPB受信機構12の処理結果を受け、PB入力か音声入力かの判定を行い、音声入力と判定した時は音声認識機構11の処理結果を採用し、PB入力と判定した時はPB受信機構12からの処理結果を採用する。選択機構14の処理結果は一括入力解釈会話機構15に渡され、第6図の場合と同様に一括入力された複数項目が正しく解釈される。

第2図は音声とPB信号が混在した一括入力方式による利用者と計算機センタの会話シーケンス例を示したもので、数字にはPB入力、それ以外には音声入力を示した例である。第3図は、この音声、PB入力を判別する選択機構14の処理を説明する図である。

号公報)。この結果、PB入力(入力2、入力3、入力5、入力6)は、PB受信機構12からT<sub>p</sub>の処理遅延後、処理結果(コード)P<sub>2</sub>、P<sub>3</sub>、P<sub>5</sub>、P<sub>6</sub>が必ず出力され、音声を誤ってPB受信結果として出力することはない。

このように、周波数フィルタにより音声でPB受信機構12が動作しないようにできても、音声認識機構11がPBを誤って認識する危険性があり、PB入力が音声認識機構11、PB受信機構12の両方から出力されることがある(入力2のケース)。さらには、音声認識機構11の処理遅延T<sub>v</sub>、PB受信機構12の処理遅延T<sub>p</sub>の違いにより、入力順序と処理結果の出力順序が逆転することがある(入力4と入力5のケース)。このため、選択機構14では以下の処理を行う。なお、説明は第3図に示したようにT<sub>v</sub>>T<sub>p</sub>として行うが、逆の場合でも全く同様である。

処理遅延の小さいPB受信機構12の出力が実際の出力時刻より処理遅延の差|T<sub>v</sub>-T<sub>p</sub>|だけ遅れたと見なし、処理遅延の差を補正する。具

体的には、PB受信機構12の出力を遅延回路等により「T<sub>v</sub>-T<sub>p</sub>」遅らせて出力する方法、PB受信機構12、音声認識機構11からの出力時刻に処理遅延時間を加減することにより、計算上行う方法などが考えられる。

選択機構14では、上記処理時間の差を補正したPB受信機構12、音声認識機構11の処理結果に対し、両者の時間差がある値dより小さかった場合には、PB入力がPB受信機構12、音声認識機構11の両者で処理されたものとみなし、音声認識機構11からの出力を破棄する。dよりも大きかった場合には、遅延時間の差を補正した結果をそのまま出力する。なお、dは利用者が音声またはPBを入力する時間間隔T<sub>i</sub>に比べ小さいため、利用者が独立に入力した音声とPBを誤って同一入力とみなし捨ててしまうことはない。この結果、①PBが音声認識機構11で誤って処理された時の出力を削除することが出来る(入力2のケース)。②処理遅延の差により出力が逆転した時にも、元の正しい順序に戻すことができる

(入力4、入力5のケース)。

以上により、音声入力に対しては音声認識機構11の処理結果、PB信号に対してはPB受信機構12の処理結果を必ず得ることができる。選択機構14の処理結果は第6図における音声認識機構11の出力と同じであり、これ以降は一括入力解釈会話機構15により、従来と同様に一括入力された複数項目が正しく解釈される。

〔発明の効果〕

以上述べたように、本発明によれば、音声とPBを組み合わせた複数項目の一括入力が可能となるので、従来技術に比べ以下の効果がある。

- ① 音声またはPBによるガイダンス入力と比較し、入力インタラクション数を減少できる。
- ② 音声あるいはPBのみによる一括入力方式と比較し、数字をPB入力、数字以外を音声入力とすることにより、

イ、数字の入力が確実になる。

ロ、電話から受発注を行うような場合、商品名(非数字単語の音声入力)と個数(数字のPB入

力)のペアで構成される入力項目が複数一括入力されるが、このようなケースでは、確実なPB入力による数字が入力項目の区切りとできるので、一括入力の解釈を確実なものとするができる。ハ、数字の音声認識用の標準ボタンが不要になるので、認識装置の価格を安くすることができ、または、その分、他の認識単語数を増すことができる。

#### 4. 図面の簡単な説明

第1図は本発明の一実施例の構成図、第2図は音声とPB信号が混在する会話シーケンス例を示す図、第3図は第1図の選択機構の処理を説明する図、第4図は従来のガイダンス方式の構成例を示す図、第5図は第4図に対する会話シーケンス例を示す図、第6図は従来の一括入力方式の構成例を示す図、第7図は第6図に対する会話シーケンス例を示す図である。

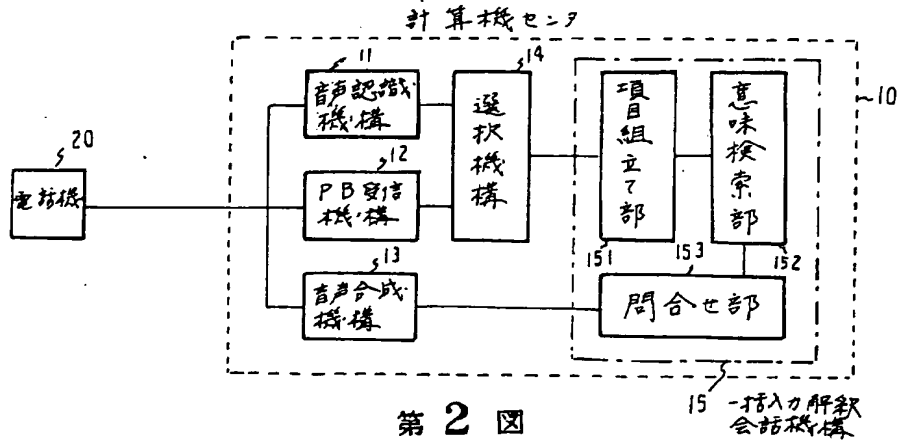
- 10…計算機センタ、 11…音声認識機構、
- 12…PB受信機構、 13…音声合成機構、
- 14…選択機構、

15…一括入力解釈会話機構、 20…電話機。

代理人弁理士 鈴木



第 1 図



第 2 図

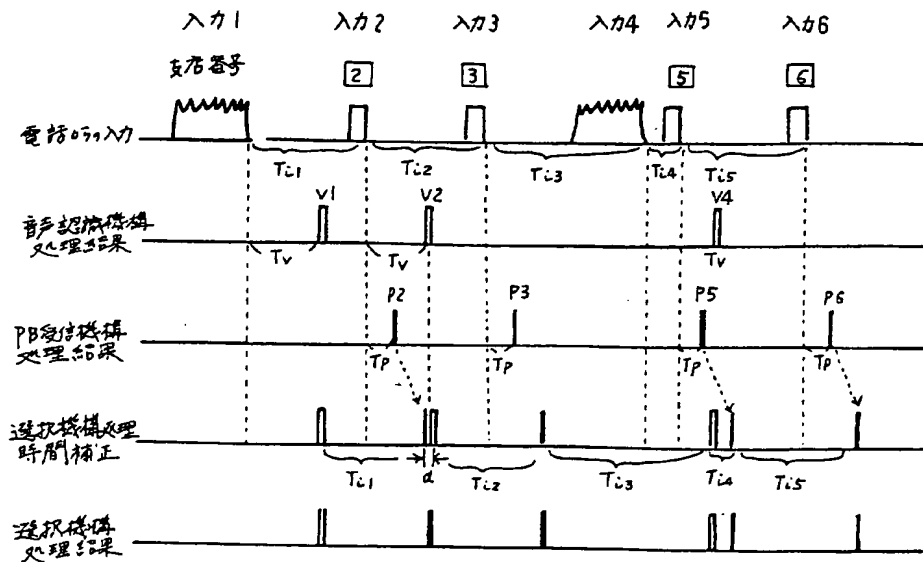
《利用者》

《計算機センタ》

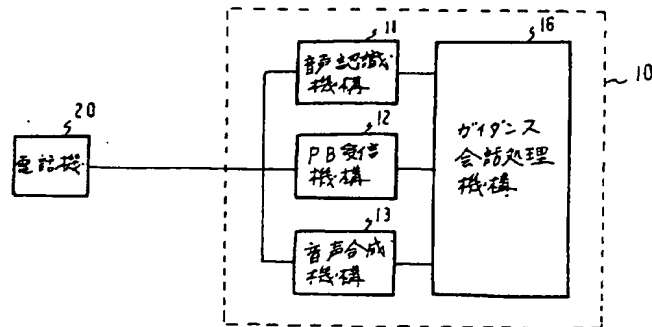
支店番号 2 3 □座番 5 6 …照会 →

← 支店番号 2, 3 □座番 5, 6 …の照会  
で済む。

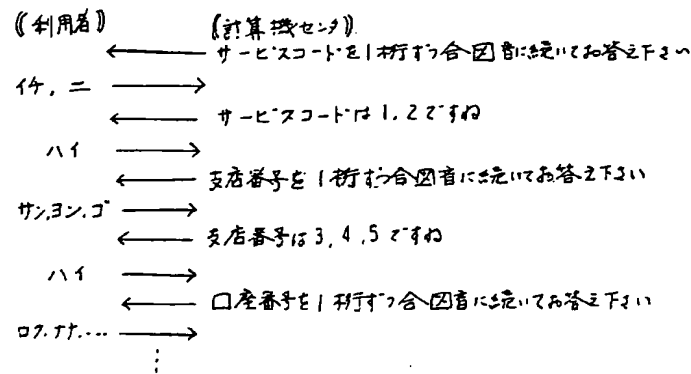
第 3 図



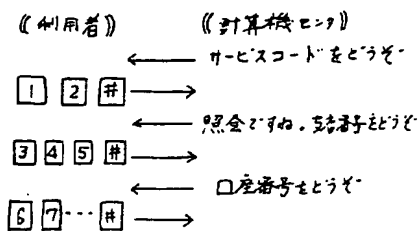
第 4 図



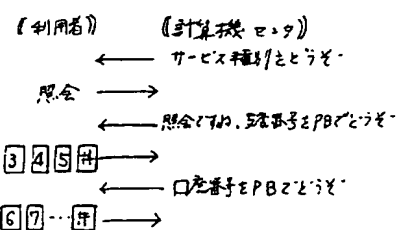
第 5 図  
(イ) (音声入力)



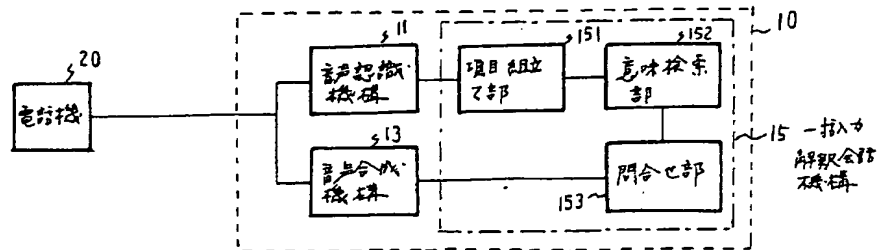
(ロ) (PB入力)



(ハ) (音声、PB入力)



第 6 図



第 7 図

